Genome-wide analysis of the regulation of sugar transport in the archaeon Sulfolobus solfataricus

Lubelska, Joanna Mariola

Chapter 3

Characterization of a deletion mutant of AraV, the ATPase of the arabinose transporter in the hyperthermophilic archaeon *Sulfolobus solfataricus*

Joanna M. Lubelska, Marleen van Wolferen, Sonja-V. Albers, Jacek Lubelski[+], Oscar P. Kuipers[+] and Arnold J.M. Driessen

[+] Department of Genetics, Rijksuniversiteit Groningen, Kerklaan 30, 9751 NN Haren

Abstract

Arabinose uptake in the hyperthermophilic archaeon *Sulfolobus solfataricus* occurs via a high-affinity arabinose binding protein dependent ABC transporter. Proteins involved in arabinose metabolism and transport are induced by the presence of arabinose in the growth medium. AraV is the ATP binding component of the arabinose transporter which contains an additional carboxyl-terminal domain that is homologous to a region in MalK, the ATP binding subunit of the maltose ABC transporter of *E. coli* that is known to interact with MalT, the transcriptional activator of the *mal* operon. To identify possible AraV-interacting regulatory protein(s), we have constructed a *S. solfataricus* strain lacking *araV.* The strain shows a delayed growth on arabinose as sole carbon source but growth is unaffected on glucose-containing medium. These data suggest the presence of an alternative transport pathway for arabinose. DNA microarray analysis revealed the upregulation of two DNA-binding proteins and of a transporter belonging to the major facilitator superfamily.

Introduction

Archaea are phylogenetically separated from bacteria based on small subunit rRNAs as a distinct group of organisms (Woese and Fox, 1977;Woese *et al.*, 1990). Their cellular organization resembles that of bacteria, but molecular details of archaeal processes such as replication, transcription and translation are similar to those known from eukaryotic organisms (Allers and Mevarech, 2005;Aravind *et al.*, 2005;Barry and Bell, 2006;Bell and Jackson, 2001). Interestingly, most of the described archaeal regulatory proteins are homologous to bacterial proteins (Aravind and Koonin, 1999).

Sulfolobus solfataricus is a hyperthermophilic crenarchaeon thriving at temperatures of 80 ºC and pH values between 2 to 4. Recently, several features characteristic for catabolite repression have been reported for *Sulfolobus* (Haseltine *et al.*, 1996;Haseltine *et al.*, 1999a;Lubelska *et al.*, 2006). Sugar and amino acid metabolism are linked, although the regulatory factors involved in this process have not been identified up to date. A factor called *car* affects the expression of β-glycosidase, α-amylase and β-galactosidase, but the exact regulatory mechanisms are unknown (Haseltine *et al.*, 1999b;Hoang *et al.*, 2004).

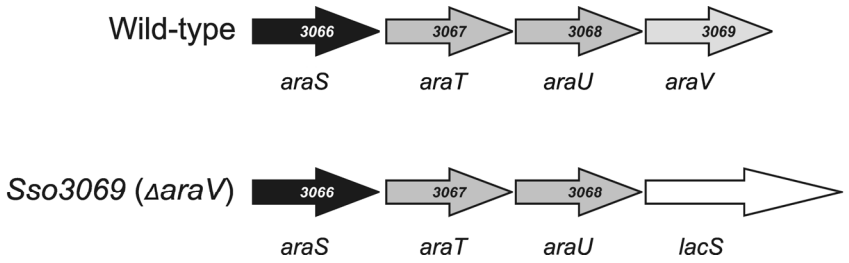

Fig.1. Genetic analysis of the *ara* cluster genes in the wild-type and *Sso3069::lacS* mutant strain.

In archaea, sugars can enter the cell via primary (ABC-type) or secondary transporters (Albers *et al.*, 2001;Koning *et al.*, 2002). The expression of the genes encoding the arabinose ABC transporter in *S. solfataricus* is induced by the presence of the substrate, D-arabinose, in the growth medium (Lubelska *et al.*, 2006). The genes of the arabinose transporter are arranged in an operon (Fig. 1A) which consist of four genes: *Sso3066* (*araS*) coding for a membrane anchored sugar–binding protein, *Sso3067* (*araT*) and *Sso3068* (*araU*) encoding two membrane domains, and *Sso3069* (*araV*), a cytoplasmic ATPase (Elferink *et al.*, 2001). These four genes are transcribed in the same direction (Elferink *et al.*, 2001), and the operon contains two potential promoter regions, one upstream of *araS* and one upstream of *araT*. Only the presence of the *araS* promoter has been confirmed experimentally (Lubelska *et al.*, 2006).

AraV is highly similar to GlcV, the ATPase of the ABC glucose transporter of *S. solfataricus* (Albers *et al.*, 1999). The crystal structure of GlcV shows the presence of two distinct domains: an N-terminal ATPase domain and C-terminal domain with an OB-like fold of unknown function (Verdon *et al.*, 2003). AraV and GlcV show sequence and structural homology with MalK, the ATPase of the maltose ABC transporter of *E. coli*, *Thermococcus litoralis* and *Pyrococcus furiosus* (Chen *et al.*, 2003;Bohm *et al.*, 2002;Lee *et al.*, 2003). In the *E. coli* MalK, the carboxyl-terminal domain functions as a binding region for the regulatory protein, MalT. MalT functions as an activator controlling the expression of the *mal* operon genes in a process known as inducer exclusion (Schreiber and Richet, 1999;Panagiotidis *et al.*, 1998). MalT acts as an activator only in the presence of maltotriose and ATP; whereupon it induces the expression of the *mal* genes. When there is no substrate present in the medium, MalT remains bound to the maltose transporter via MalK and thus gene transcription of *mal* operon is not activated (Boos and Shuman, 1998;Boos and Bohm, 2000). In *malK* deletion mutants the *mal* operon is constitutively expressed because of lack of the interacting partner for MalT (Bukau *et al.*, 1986). Despite the structural homology, so far no homologs of MalT have been identified in *T. litoralis* and *P. furiosus* that could potentially interact with MalK. Instead, in *Pyrococcus* another regulator has been identified, i.e., TrmB which represses the genes encoding the ABC transporter for trehalose and maltose (the TM system) and for maltodextrins (the MD system) (Lee *et al.*, 2003). Interestingly, the TM system is induced by trehalose and maltose, while the MD system uses

maltotriose and sucrose as inducers (Lee *et al.*, 2003). Currently, there is no evidence for a direct interaction between TrmB and MalK, and therefore the role of the regulatory domain in the archaeal transporter ATPases remains to be elucidated (Lee *et al.*, 2003). Because of the presence of the putative regulatory domain in the *S. solfataricus* AraV and GlcV, one might expect a regulatory mechanism of transport activity with similarity to that found for the maltose system in *E. coli*. However, by database searches no homologs of MalT or TrmB could be found in the *S. solfataricus* genome. To identify possible regulatory proteins controlling the expression of the arabinose ABC transporter genes, and thus potential binding partners for the regulatory domain of AraV, we introduced a targeted gene deletion of *Sso3069* (*araV*) gene. Growth and uptake studies were performed to characterize this strain in comparison to the wild-type and DNA microarray analysis was used to identify differentially expressed genes leading to the identification of two putative regulators of sugar transport and metabolism in *S. solfataricus*.

Materials and methods

Strains and culture conditions

Sulfolobus solfataricus cells were grown aerobically at 80 ^{o}C in defined medium (Brock *et al.*, 1972) supplemented either with 0.05% tryptone (w/v) and 0.4% D-arabinose or D-glucose (for binding and uptake assays), or sugars alone (growth curves) as a sole carbon and energy source. Medium was adjusted to pH of 3 with sulphuric acid. Growth was monitored spectrophotometrically at 600 nm.

Construction of knockout strain

S. solfataricus PBL2025, the wild-type in this study (Haseltine *et al.*, 1999b) was used to construct the PBL2025 *Sso3069::lacS* mutant strain. pET2268, a *lacS* containing plasmid, was used for the gene disruption essentially as described before (Szabo *et al.*, 2007;Worthington *et al.*, 2003;Albers *et al.*, 2008)**.** Primers design was based on the genome sequence of *S. solfataricus* P2 (http://www-archbac.u-psud.fr/projects/sulfolobus). The primers used correspond to approximately 400 base pairs of the flanking regions of the *Sso3069* gene, i.e., for the downstream region: forward23 5'-ccccccATGGGTACCATCATGTGGTTGGTTCCGTTGTACGCTATG-3' (KpnI) and reverse24 5'- cccccc<u>CCATGG</u>TTACTCTACACCCCTTTAGCTCCTCCACC-3' (NcoI); and the upstream region: forward25 5'-cccccc<u>ATGGGATCC</u>GTTGGGGAAGTTTTATATTTATTAAAACATTAT-3' (BamHI) and reverse26 5'-cccccc<u>GCGGCCGC</u>TTAACCTAGTTGGTCCTCCCTCAGCACCAACGA-3' (NotI) Flanking regions were cloned into pET2268 resulting in the pET2268-araVflanks construct.

Southern blotting

Genomic DNA (10 µg) was digested with the appropriate enzymes and separated on 0.8% agarose gel. The gel was equilibrated in 20x SSC buffer (3 M sodium chloride; 0.3 M sodium citrate) and the DNA was transferred overnight to a Zeta-probe positively charged membrane (BioRad). DNA hybridization was performed in standard hybridization buffer. Both *lacS* and *Sso3069* probes were generated by PCR reaction and labeled with digoxigenin using HighPrimeKit (Roche, The Netherlands). Blots were exposed to anti-DIG conjugated with alkaline phosphatase antibodies (Sigma) and developed by chemiluminescence using CDP-star (Roche Applied Science) and visualized on a Lumi Imager (Roche Applied Science).

Western blotting

Western blot analysis was performed as described (Lubelska *et al.*, 2006) using antibodies raised in chicken for GlcS (*Sso2847*) and AraV (*Sso3069*) or in rabbit for GlcV (*Sso2850*).

Purification of binding proteins

Membranes isolated by high-speed centrifugation were resuspended in 10 ml of buffer A (20 mM MES, pH 7.5, 100 mM NaCl and 1% Triton X-100) and incubated at room temperature for one hour. Solubilized membrane proteins were subjected to a ConA sepharose column equilibrated with 10 volumes of buffer A. Additionally, the flow through was reloaded twice. The column was washed with 30 volumes of buffer A, and bound membrane proteins were eluted with buffer B (buffer A supplemented with 200 mM methyl-α-D-mannopyranoside). Eluted fractions were analyzed on 10% SDS-PAGE..

Sugar binding and uptake assays

S. solfataricus cells were collected when they reached the logarithmic phase at OD_{600} of 0.6. The cells were washed in medium without addition of a carbon source and resuspended to a concentration of 10 mg protein/ml. For the binding assays, binding proteins (ConA fractions) were isolated as described above. The pH was adjusted to 3 with sulphuric acid. An aliquot of 10 µl of cells or the ConA fraction was added to 90 µl of a pre-warmed minimal Brock medium. Radiolabeled sugar (1 µl), either D-[^{14}C]-glucose or D-[^{14}C]-arabinose (Amersham and ARC Inc.) was added. The mixtures were incubated at 60 °C for 2 min and the reactions were stopped by adding 2 ml of ice-cold 0.1 M lithium chloride, and rapidly filtered through 0.45 µm nitrocellulose filters BA 85 (PROTRAN, Schleicher & Schuell). Filters were dissolved with scintillation fluid (Packard) and the radioactivity was measured by scintillation counting (FLOscint A, Packard Instruments). The data presented is an average of three independent measurements.

Expression of SSO3061 and SSO3188 in Sulfolobus solfataricus M16

For expression of SSO3188 and SSO3061 in *S. solfataricus* (Albers *et al.,* 2006b), a culture of strain M16 (*pyrEF/lacS* double mutant) was grown in Brock medium supplemented with 0.1% NZ-amine and 10 µg/ml uracil to an OD_{600} of 0.2-0.3. Cultures were cooled on ice and spun down at 3500 x g for 15 min at 4 ºC (Allegra 6R centrifuge, Beckman Coulter). The obtained pellet was resuspended and centrifuged 3 times with 50, 25 and 1 ml of cold 20 mM sucrose solution, respectively. Washed cells were then adjusted to 10^{10} cells/ml in the sucrose solution. The virus vectors pSVA171 or pSVA172 (300 ng) (dialyzed against MilliQ) (Table 1) were mixed with 50 µl of competent cells and transferred to a 0.2 cm Gene Pulser Cuvette (BioRad). Directly after electroporation (1.5 kV, 25 µF, 400 Ω), 1 ml of medium was added. Cells were regenerated at 75 ºC for 1 hour and subsequently inoculated into pre-warmed 50 ml cultures with uracil. After the cells reached an OD_{600} of 0.5, 1 ml of the culture was transferred to 50 ml selective medium without uracil. As soon as the cells started to grow (OD_{600} ~ 0.2), 10 ml of cell culture was diluted into 400 ml of the same medium. Cells were grown to an OD_{600} of about 0.5; during cultivation pH of 3 was maintained by adjusting regularly with sulfuric acid. The temperature was subsequently elevated to 84°C for 24 h, whereupon cells were harvested by centrifuging for 15 min at 3500 x g (Allegra 6R centrifuge, Beckman Coulter) and resuspended in TE (10 mM Tris, 1 mM EDTA, pH 8.0). Cells were lysed using the MSE Soniprep 150 (BIOCON Scientific). Sonication was performed for 10 cycles of 15 sec sonication and 45 sec vortexing. After sonication, each sample was centrifuged at 3000 x g for 10 min to pellet the remaining unlysed cells, supernatant was cleared by ultracentrifugation for 20 min at 80000 x g (Optima™ Ultracentrifuge, Beckman Coulter). Obtained cytoplasm (supernatant) and membranes (pellet) were used for SDS-PAGE and western immunoblotting as described in Western immunoblot analysis.

Expression of SSO3061 and SSO3188 in Sulfolobus acidocaldarius

For expression of SSO3061 and SSO3188, *S. acidocaldarius* Mr31 cells were electroporated with methylated pSVA154 and pSVA155 (Table 2) with a Genepulser II (1.25 kV, 25 µF, 1000 Ω, BioRad) and 2-mm cuvettes (Thermotron). Cells were regenerated in recovery solution (1% sucrose, 20 mM beta-alanine malate buffer, pH 4.5, 10 mM MgSO4) for 30 min at 70 ºC and subsequently inoculated into pre-warmed 50 ml cultures. Once the cells reached an OD_{600} of 0.5, 1.5 ml of culture was transferred to 50 ml medium and directly induced with 0.2% maltose. Cells were grown to an OD_{600} around 0.7 and harvested by centrifugation of the cultures for 15 min at 3000 x g (Allegra 6R centrifuge, Beckman Coulter). Pellets were resuspended in 2 ml TE (10 mM Tris, 1 mM EDTA, pH 8.0), cell suspensions were then lysed using the MSE Soniprep 150 (BIOCON Scientific). Sonication was performed for 10 cycles of 15 sec sonication and 45 sec vortexing. After sonication, each sample was centrifuged at 3000 x g for 10 min to pellet the

remaining unlysed cells. To obtain the cytoplasm of the cells, the supernatant was centrifuged for 30 min at 80000 x g (Optima™ Ultracentrifuge, Beckman Coulter).

Preparation and refolding of inclusion bodies

E. coli Tuner cells (Novagen) were transformed with pSVA167 (Table 1) via a heat shock at 42 ºC. Bacteria were grown to an OD_{600} ~ 0.6 and placed on ice for 30 min. Expression of SSO3061 was induced by adding 0.5 mM IPTG. Growth was continued for another 4 hours; cells were then harvested by centrifugation for 30 min at 6000 x g (JLA 8.100). Cell pellets were resuspended in lysing buffer (50 mM Tris-HCl pH 8.0, 100 mM NaCl, 5 mM EDTA, 0.1% NaN_3, 0.5% Triton-X100, 0.1 mM PMSF and 1 mM DTT) and lysed using the MSE Soniprep 150 (BIOCON Scientific). Sonication was performed for 15 cycles of 15 sec sonication and 45 sec vortexing. After sonication, 10 mM $MgSO_4$ was added to chelate the EDTA. The lysate was supplemented with some DNase and lysozyme and incubated at room temperature for 20 min. To collect inclusion bodies the suspension was centrifugated for 30 min at 8000 x g. The pellet was dispersed by sonication and washed twice in the lysing buffer without DNase and lysozyme and once in the lysing buffer without DNase, lysozyme and Triton-X 100. After another centrifugation step, washed inclusion bodies were dissolved in 6 M GuHCl, 0.5 M NaCl and 20 mM Hepes, pH 8.0. Obtained proteins were applied to 1 ml Ni-NTA beads (Sigma) (pre-equilibrated in the dissolving buffer) and incubated with constant turning at 4 ºC for one hour. The protein-resin mixture was then transferred to an empty column. The column was washed 10 times with dissolving buffer without imidazole, and 10 times with buffer containing 10 mM imidazole. The bound proteins were eluted with 300 mM Imidazole at a lower pH of 7.0. For SDS-PAGE 30 µl of the different purification fractions was TCA precipitated and brought on a 15% SDS PAGE.

Refolding was done with a number of overnight dialyzing steps. First 10 ml of 250 µg/ml purified and denaturated protein (elution fractions in 6M GuHCl, 0.5 M NaCl and 20 mM Hepes, pH 6.0) was dialyzed against 0.5 M arginine, 2M GuHCl, 50 mM Tris, pH 6.0. Then pH was lowered to 3.5 with 20 mM Na-Acetate, 50 mM KCl, 1M GuHCl, pH 3.5. Finally proteins were brought to a neutral pH of 6.5 by dialysis against 20 mM MES, 50 mM KCl. Precipitant was removed by 5 min of centrifugation in a table top centrifuge.

Electrophoretic mobility shift assay (EMSA)

The 273 bp promoter region of *araV* was amplified using cy3 labeled primers designed according to the genomic sequence of *S. solfataricus* P2: 5'-[Cy3]GAGATGAAGCTTAGAAGATC-3' and 5'-[Cy3]CATATTCTCGGGTACTTTTATGACC-3'. PCR product was then purified using the GFX PCR DNA and Gel Band Purification Kit (GE Healthcare). The binding reaction was performed for 15 min at 60 °C in binding buffer (1x SBA), with 50 nM labeled PCR product and 6 µM refolded

SSO3061 in a total volume of 12 µl. Samples were loaded with 3 µl of 60% glycerol onto a 5% acrylamide gel, and electrophoresed in SBA buffer, pH 9.5 for 90 min at 120 V. Labeled DNA fragments were visualized with a FUJI LAS4000 imager (Fujifilm).

DNA microarray analysis

DNA microarray experiments were essentially performed as described (van Hijum *et al.*, 2005;den Hengst *et al.*, 2005). Briefly, RNA was isolated from two biological replicate cultures of *S. solfataricus* PBL2025 (wild-type) and the two *S. solfataricus* PBL2025 *Sso3069::lacS* (mutant) strains. Strains were cultivated in minimal medium with addition of 0.4% D-arabinose and harvested at OD_{600} of 0.4. RNA was isolated by means of high pure RNA isolation kit (Roche). Subsequently, single-strand reverse transcription (amplification) and indirect labeling of 20 µg of total RNA, with either Cy3 or Cy5 dye, was performed (including a sample in which the dyes were swapped to correct for dye-specific effects). Hereafter, mixed labeled cDNA samples were overnight hybridized to microarray slides of *S. solfataricus* (Isogen) spotted in duplicates as described. Following hybridization, slides were washed for 5 min at 37 °C in 2xSSC (Standard Saline Citrate), 0.5% SDS and 5 min in 1xSSC and 0.25% SDS to remove non-specifically hybridized cDNAs. Slides were scanned using a GenePix 4200AL (Westburg). Next, individual spot intensities were measured using GenePix Pro 6.0. Slide data were processed and normalized using MicroPrep (van Hijum *et al.*, 2003), which yielded average ratio of gene expression with mutant over the wild-type strain. Expression of a gene was considered to be significantly altered when the Cyber T baysian *p* value ≤1E-05. All DNA microarray data obtained in this study are available at http://molgen.biol.rug.nl/publication/araV_sulfolobus_data/.

Results

Targeted deletion of araV encoding the ATPase of the ABC arabinose transport

For the construction of the *Sso3069::lacS* deletion strain PBL2025, a lactose auxotrophic *S. solfataricus* strain (Schelert *et al.*, 2004) was used. After the electroporation with the pET2268araV-flanks construct (Fig. 1), integrants were selected on minimal lactose medium. After plating singe blue colonies were picked and analyzed. The selection of the *Sso3069::lacS* (*ΔaraV*) strain was confirmed by PCR (data not shown) and Southern (Fig. 2) analyses.

Characterization of the Sso3069::lacS (ΔaraV) deletion mutant

Growth features. Growth of wild-type and *ΔaraV* cells was compared in minimal Brock media supplemented with either tryptone, D-glucose or D-arabinose. Whereas both strains grew equally well on glucose (Fig. 3A), on D-arabinose the knockout strain showed a 24 hours prolonged lag

phase in comparison with the wild-type (Fig. 3B). It seems likely that during the lag phase, the cells activate an alternative transporter that allows cellular entry of arabinose.

Since fructose and xylose were shown to be competitors for arabinose binding to the arabinose substrate binding protein (Elferink *et al.*, 2001), growth on these sugars was also tested. However, both the wild-type and knockout strain grew poorly on fructose, whereas xylose did not support growth at all (data not shown). The growth studies on fructose and xylose were previously performed with the P2 strain. Our data indicates that none of these sugars can be used as a carbon source by the PBL2025 strain.

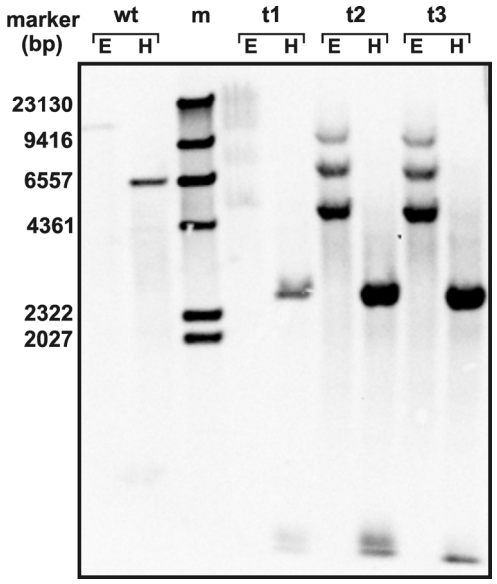

Fig.2. DNA analyses of a wild-type and *Sso3069::lacS* strains. Southern blot with the genetic materials restricted with either EcoRI (E) or HindIII (H) and hybridized with probes against *lacS* and araV. wt – wild-type, after restriction with EcoRI gives a band of 12776 base pair, with HindIII – 6600 bp; single transformants, marked as t1, t2, t3, EcoRI: 7600, 5100, 500 and HindIII 4100, 2650, 700. Single transformant named t3 was used in the further analysis in our studies.

Expression levels of the arabinose and glucose ABC transporter. In order to confirm the deletion of *araV* and check the influence of the knock out on the other genes in the same operon protein levels have been checked in western blot analyses. The presence of arabinose in the medium induces the expression of the arabinose transporter (Lubelska *et al.*, 2006) and enzymes involved in arabinose metabolism (Brouns *et al.*, 2006). As expected *araV* is expressed in the wild-type when grown on arabinose, whereas no AraV protein was detected in the *ΔaraV* strain. This result confirms the deletion of *araV*. On the other hand, the protein levels of AraS were comparable in wild-type and *ΔaraV* cells (Fig. 4A and B), while AraS and AraV are not expressed in media containing tryptone or glucose.

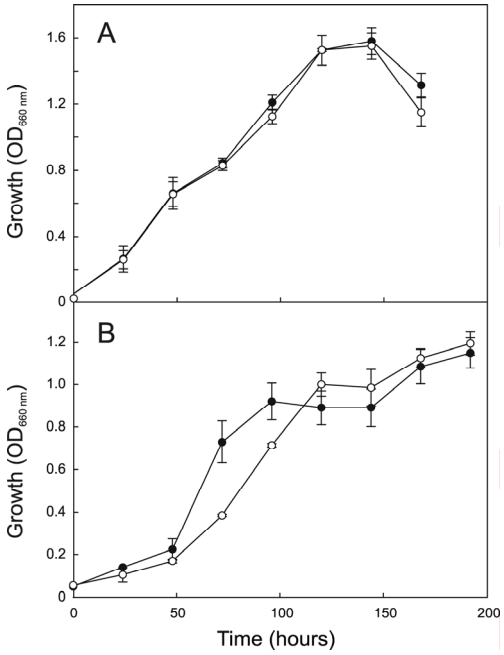

Fig.3. Growth patterns of the PBL2025 (filled symbols) and *Sso3069::lacS* (open symbols) on A. glucose and B. arabinose.

The glucose ABC transporter is expressed constitutively (Lubelska *et al.*, 2006). As a control, the levels of GlcS, a glucose-binding protein, and GlcV, an ATPase, were analyzed by western blotting in the *ΔaraV* and the wild-type, and found to be similar under all tested growth conditions in both the wild type and araV deletion strain (Fig. 4C).

Binding experiments. The sugar binding activities were determined by means of radiolabeled arabinose and glucose using ConA fractions isolated from the membranes derived from wild-type and *ΔaraV* cells. The specificity of glucose or arabinose binding was confirmed by means of competition experiments using an excess of cold arabinose (grey bars) or glucose (white bars), respectively (Fig. 5). Comparable levels of glucose binding were observed with the ConA fractions isolated from wild-type and mutant strains, either grown on glucose or arabinose and the level of binding was in the range of 30-40 nmoles / mg of total proteins (Fig. 5A, black bars). An excess (1 mM) of unlabeled glucose chased the bound radioactively labeled glucose, while arabinose was found to compete very poorly with glucose for binding to GlcS.

Significant arabinose binding was detected in the ConA fractions (Fig. 5B, black bars) isolated from cells grown on arabinose, both with the wild-type and mutant cells. An excess of unlabeled arabinose (1 mM, 1000-fold higher than the radiolabeled sugar) effectively chased away the AraS-bound arabinose. On the other hand, unlabeled glucose weakly completed with arabinose for binding to the AraS protein (Fig. 5B).

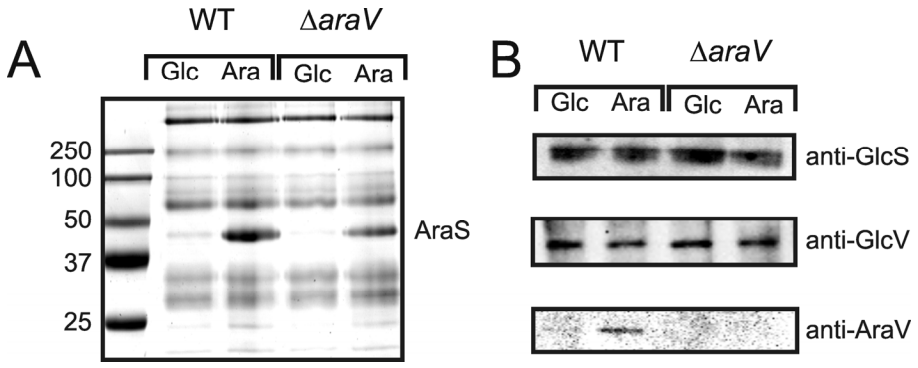

Fig.4. Expression levels of the glucose and arabinose ABC transporter subunits. A. Protein profiles of ConA fractions isolated from wild-type and *Sso3069::lacS* membranes derived from cells grown in the presence of glucose (Glc) or arabinose (Ara). B. Western blot analyses of protein expression levels in the wild-type and deletion strains of GlsS (ConA fraction) and GlcV and AraV (membranes). Double band in of GlcS western blot is the result of proteolysis.

Uptake studies. Cells were also assayed for the uptake of arabinose and glucose. As described above, the glucose ABC transporter genes are not up/down-regulated in the mutant, so as expected both wild-type and deletion mutant cells showed similar rates of glucose uptake, irrespective as to whether they were grown on glucose or arabinose. On the other hand, attempts to observe arabinose uptake by whole cells were unsuccessful.

DNA microarrays. DNA microarrays were employed to identify the most differentially expressed genes in the *S. solfataricus Sso3069::lacS* mutant strain as compared to the wild-type when grown on arabinose. Table 1 shows the folds of up- or downregulation of the statistically significant differentially expressed genes. *Sso3019*, coding for the β-galactosidase, was the highest up-regulated gene. Other up-regulated genes were *Sso3188*, coding for a putative transcriptional regulator, *Sso3003* – a glucose-1-dehydrogenase, *Sso0091* – a ribosomal protein, and two hypothetical proteins *Sso12127* and *Sso2745, respectively*. Expression of *Sso3118*, encoding KdaD, an enzyme involved in catabolism of D-arabinose was also up-regulated. *Sso2113* was the only upregulated gene coding for a putative transport protein.

The down-regulated genes were subunits of the arabinose ABC transporter: *Sso3066* – AraS and *Sso3067* – AraT and also genes located in close proximity to the *ara* genes, such as *Sso3061* which encodes a putative transcriptional regulator and *Sso3053* – the maltose binding protein, MalE (Elferink et al., 2001). *Sso3124*, the gene encoding AraD involved in arabinose utilization is also down-regulated. Down-regulated was also a group of genes encoding for enzymes involved in biosynthesis of lysine, such as *Sso0159* (LysX) *Sso0160* (LysJ) and *Sso5317* (a homologue of LysW).

Concluding, DNA microarray data revealed 2 potential regulators: *SSO3061* and *SSO3188* are predicted DNA-binding proteins possibly controlling the expression of the arabinose transporter genes and other genes involved in sugar metabolism. Also, the arrays showed the upregulation of other transporter gene and possible a link with the regulation of the maltose utilization genes. Moreover, deletion of *araV* influenced expression of the genes linked to amino acid metabolism.

Table.1. DNA microarray analysis showing differentially expressed genes.

Genes up-regulated		
Gene	**Expression ratio**	**Description**
SSO3019	3.06	LacS, β-galactosidase
SSO3188	1.69	Hypothetical protein, predicted transcriptional regulator
SSO3003	1.64	(dhg-1) Glucose 1-dehydrogenase
SSO0091	1.61	(rpl7AE) LSU ribosomal protein L7AE
SSO12127	1.52	Hypothetical protein
SSO2745	1.46	Hypothetical protein
SSO3118	1.45	2-keto-3-deoxy-D-arabinonate dehydratase (KdaD)
SSO2113	1.44	Metabolite permease, putative
SSO2323	1.40	Flagellin (FlaB)
SSO2629	1.40	Oxidoreductase (flavoprotein)

Genes down-regulated		
Gene	**Expression ratio**	**Description**
SSO0127	2.88	Putative (leuA-1) 2-isopropylmalate synthase
SSO3067	2.53	Arabinose ABC transporter, permease, AraT
SSO3066	2.32	Arabinose ABC transporter, arabinose binding protein, AraS
SSO3053	2.25	Maltose ABC transporter, maltose binding protein
SSO1578	2.20	Putative DsrP protein
SSO0159	2.01	Ribosomal protein S6 modification protein (LysX)
SSO0160	1.91	Acetylornithine aminotransferase (LysJ)
SSO2797	1.80	Conserved hypothetical protein
SSO0977	1.80	(leuA-2) 2-isopropylmalate synthase
SSO1816	1.77	Hypothetical protein
SSO8566	1.72	Hypothetical protein
SSO5317	1.71	Protein with homology to LysW *Aeropyrum pernix* K1
SSO3124	1.67	Mandelate racemase /muconate lactonizing enzyme related protein (AraD)
SSO6845	1.67	Hypothetical protein
SSO3061	1.64	DNA binding protein with HTH motif
SSO0538	1.63	Hypothetical protein
SSO2285	1.59	Hypothetical protein predicted ATPase
SSO2043	1.59	Amino acid transporter related protein
SSO1641	1.57	Conserved hypothetical protein
SSO3139	1.51	Conserved hypothetical protein : predicted inorganic ion transport and metabolism
SSO1170	1.50	Predicted ABC-type sugar transport systems, permease

Expression studies of the potential regulators SSO3061 and SSO3188.

Since the microarray studies indicated that SSO3061 and SSO3188 might play a role in the regulation of the arabinose transporter operon, both proteins were overexpressed in *E. coli*. Unfortunately, all trials to express SSO3188 in different available *E. coli* strains were unsuccessful as cells started lysing immediately upon the induction of expression of the protein.

Overproduction of his-tagged SSO3061 in *E. coli* resulted in inclusion body formation. Refolding procedures yielded soluble protein that was purified by Ni-NTA affinity chromatography.
To overcome the expression problems in *E. coli* also the homologous expression in *Sulfolobus* was employed. First, the virus vector system for *S. solfataricus* using a heat-stress inducible promoter system was used (Jonuscheit *et al.*, 2003). Secondly, a new plasmid based vector system for *S. acidocaldarius* was employed (S. Berkner, unpublished). However, both methods failed to produce any detectable amounts of overexpressed proteins.
Refolded SSO3061 purified from *E. coli* was used in electrophoretic mobility shift assays (EMSA) using the cy3-labeled *araS* promoter region (273 bps). A clear protein specific shift of the *araS* promoter region was observed upon the addition of refolded SSO3061 (Fig 6, lanes 2-5). However, by addition of unspecific competitor DNA the shift disappeared (Fig. 6, lane 6). More detailed EMSA experiments confirmed that the observed up shift in presence of SSO3061 was independent of the identity of the added DNA.

Discussion

Uptake of arabinose and glucose by *S. solfataricus* is mediated by binding protein-dependent ABC transporters. The ATPase subunits of these transporters, i.e., AraV and GlcV, respectively, show high sequence and structural similarity to MalK, the ATPase of the maltose transporter of *E. coli* (Chen *et al.*, 2003) and to homologs of the archaea *P. furiosus* (Lee *et al.*, 2005) and *T. litoralis* (Bohm *et al.*, 2002). These subunits have an unusual structure: in addition to the conserved N-terminal ATP-binding (ABC) domain, they possess a C-terminal extension with a barrel-like structure (Chen *et al.*, 2003;Verdon *et al.*, 2003;Diederichs *et al.*, 2000). In *E. coli*, the latter domain functions as the site of binding of the transcriptional activator MalT that functions in a process termed inducer exclusion (Bohm *et al.*, 2002). The function of the corresponding domains of the *S. solfataricus* AraV and GlcV is unknown. In contrast to the arabinose transporter, the glucose transporter is constitutively expressed (Lubelska *et al.*, 2006).
Expression of the arabinose transporter genes is observed only when D-arabinose is added to the growth medium while they are repressed by the presence of a specific set of amino acids. To determine the possible role of AraV in the regulation of sugar metabolism and transport, we constructed an *araV* deletion strain that was characterized to some detail. Although, the growth of the *ΔaraV* strain showed a considerable lag phase, cells eventually started to grow on arabinose. A fresh transfer of these cells to a medium with arabinose allowed immediate growth suggesting the induction of an alternative transport system for arabinose.

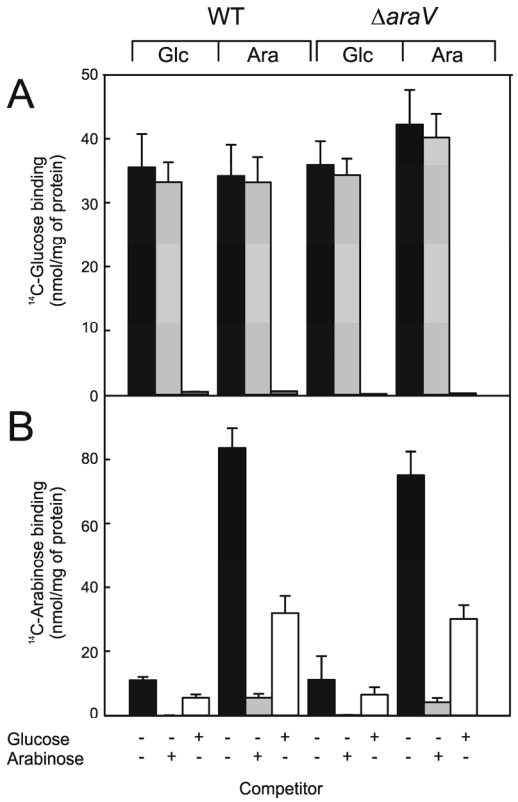

Fig.5. Binding of radiolabeled glucose (A) or arabinose (B) to the ConA fractions isolated from wild-type and Δ*araV* cells cultivated either on glucose or arabinose. Specificity of binding was measured in the presence of 1000-fold excess of a cold glucose (white bars) or arabinose (grey bars). The concentration of the radiolabeled sugar was 1 µM.

On the other hand, growth on tryptone and glucose was identical for the wild-type and knockout strain. The *ΔaraV* strain showed a slightly decreased level of arabinose binding as compared to the wild-type, but glucose binding was similar for both strains (Fig. 5). Furthermore, both wild-type and mutant grown on arabinose showed similar glucose uptake when cultivated either on glucose or arabinose. Unfortunately, the uptake activity for arabinose could not be accessed as various attempts to detect arabinose uptake were unsuccessful. Nevertheless, the ability of the *ΔaraV* strain to grow on arabinose suggests the presence of an alternative transport system, or, although less likely, a possible crosstalk between ABC domains of the sugar transporters. Since in the *ΔaraV* strain, the *araS* and *araT* genes are still expressed, a possibility is that this system combines with an ABC domain of another transporter, such as the constitutively expressed *glcV*.

A possible clue for the identity of an alternative system for arabinose uptake follows from a global transcriptome analysis of the wild-type versus mutant grown on arabinose. The *ΔaraV* strain shows the upregulation of a putative transporter, encoded by *Sso2113* that belongs to the major facilitator superfamily (Table 1). Previous DNA microarray analysis, comparing the gene expression patterns on arabinose and glucose, revealed the expression of the arabinose ABC transporter as the only active arabinose uptake system (Brouns *et al.*, 2006). Therefore, the

upregulation of *Sso2113* seems to be *ΔaraV* strain specific event, possibly as a compensatory mechanism allowing growth on arabinose. Assessment of protein and substrate binding activities showed that the arabinose binding protein level was reduced while the glucose binding protein level was not markedly changed in *ΔaraV* strain. Deletion of *araV* negatively influenced the expression of other genes of the *ara* operon, as confirmed by transcriptome analysis. These data further indicate that the compensatory mechanisms involved do not result in elevated levels of the known ABC transporters. Future experiments should address the function of *Sso2113* and reveal its substrate specificity.

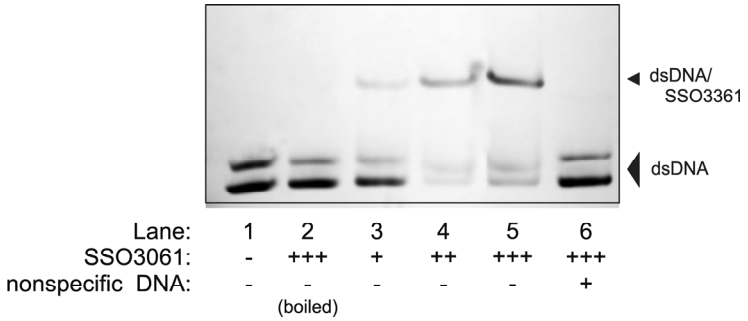

Fig.6. EMSA with an amplified cy3-labeled *ara* promoter (273 bp) region and refolded SSO3061. Binding reactions were performed with 100 ng of the cy3-labeled *ara* promoter region in binding buffer without any addition (lane 1) and with 0.225 nmol boiled SSO3061 (lane 2); 0.075 nmol refolded SSO3061 (lane 3), 0.15 nmol refolded SSO3061 (lane 4), 0.225 nmol refolded SSO3061 (lane 5), and 0.225 nmol refolded SSO3061 and nonspecific DNA (50x pET15b) (lane 6).

The gene that was most markedly upregulated in the *ΔaraV* strain was *lacS*, which encodes the β–galactosidase. This was expected as this gene was used as a marker for the knockout procedure. However, the microarray data also indicates the high level expression of *Sso3188* and the downregulation of *Sso3061*. Both genes are predicted to be transcriptional regulators according to COG classification. Interestingly, the arabinose ABC transporter is present only in *S. solfataricus*, and not in *S. acidocaldarius* or *S. tokodaii*. The two predicted regulators *Sso3188* and *Sso3061*, are also not present in these strains, which suggests that they may be directly involved in the regulation of sugar, most likely in arabinose metabolism. Expression studies only led to the expression of SSO3061 into inclusion bodies in *E. coli*. Refolded SSO3061 did bind DNA, however, the binding appeared unspecific and not restricted to the *araS* promoter region. This shows that SSO3061 is a DNA binding protein, but it remains uncertain if it is specific for arabinose metabolism. Therefore, it will be necessary to carry out a knock-out analysis to determine if these two proteins play a role in the regulation of sugar metabolism.

The gene *Sso3118* that encodes the 2-keto-3-deoxy-D-arabinonate dehydratase (KdaD), is also upregulated in the *ΔaraV* strain. On the other hand, the expression of *Sso3124* (*araD*), encoding for a mandelate racemase/muconate lactonizing enzyme related protein, is reduced in the

mutant. Both, KdaD and AraD are involved in pentose oxidation pathway of *S. solfataricus* (Brouns *et al.*, 2006). Since all these genes are upregulated in arabinose grown cells, the downregulation of *araD* expression may suggest the activation of another pathway for arabinose degradation. Likewise, the deletion of *araV*, which is located in the distal region of the *ara* operon, also resulted in the down regulation of *Sso3066* and *Sso3067*, coding for arabinose-binding protein (AraS) and permease (AraT), respectively (Table 1). However, the protein levels of AraS and the arabinose binding levels in the wild-type and deletion strain are very similar (Fig. 4). The reason for this discrepancy in mRNA and protein levels is unknown, but is encountered frequently in comparison of transcriptomic and proteomic data. Another down-regulated gene, located in close proximity of *araV* is *Sso3053*, coding for maltose-binding protein. It is well possible that one or both of the putative regulatory protein are involved in the regulation of the expression of both the *ara* and *mal* genes. A common feature of the genes involved in transport and utilization of arabinose is the presence of the 'ara-boxes' within promoter regions. The motif AACATGTT is located 10 nucleotides upstream of the TATA-box and can be found in six different promoter regions in the genome of *Sulfolobus*. Among these clusters of genes are *araS* and the genes encoding for enzymes involved in pentose oxidation pathway (Brouns *et al.*, 2006). It would be of interest to determine if one or both putative DNA-binding proteins that show an altered expression in the *araV* deletion strain bind to these promoter regions.

The genes *Sso0159* (*lysX*) and *Sso0160* (*lysJ*), encode a ribosomal protein modification enzyme and an acetylornithine aminotransferase, respectively. These genes are also down-regulated in the *ΔaraV* strain. Together with *orfF* and *argD* they belong to an operon which expression is controlled by LysM, a constitutively expressed transcriptional regulator (Brinkman *et al.*, 2002). *Sso5317* is also downregulated in *ΔaraV* strain and the corresponding protein shows sequence homology to LysW, a lysine biosynthesis protein, of *Aeropyrum pernix*. Both operons, i.e., the *lysYZM* genes that are constitutively expressed, and the *lysWXJK* genes that are induced when LysM binds to the promoter region because of a shortage of lysine, are involved in the lysine biosynthesis pathway. Since the only carbon source present in our growth experiments is arabinose, the *lysWXJK* genes are likely upregulated. Our result are consistent with the previously proposed link between sugar and amino acid metabolism (Haseltine *et al.*, 1996;Haseltine *et al.*, 1999a;Lubelska *et al.*, 2006). LysM belongs to the Lrp/AcnC family of regulators, which are widely distributed in bacteria and archaea. *E. coli* Lrp is a global regulator controlling expression of 75 genes (Brinkman *et al.*, 2003;Newman and Lin, 1995). It is very likely that LysM also plays a more general role in *Sulfolobus*, and perhaps also regulates the expression of *ara* genes. Recently, a *lysM* knock-out strain was constructed (personal communication, E. Peeters) to test this hypothesis.

Obviously, the deletion of *araV* results in a deregulation of sugar transport and metabolism. However, to determine the significance of the C-terminal domain of AraV in these regulatory events it would be of interest to test a mutant lacking only the C-terminal extension in the *ΔaraV* strain. Based on the data presented above we can speculate about a possible mechanism of *ara* gene regulation and the role of AraV in this pathway. The expression of the arabinose transporter genes might be regulated in a similar manner as the maltose/trehalose transporter of *Pyrococcus*. When arabinose is absent in the medium, a regulatory protein may remain bound to the promoter region and represses *ara* gene expression. In the presence of arabinose, the regulator may bind to AraV. This would result in the de-repression of *ara* gene expression. It is likely that such a regulator would control the expression of all 'ara-box' containing promoters. In the mutant that lacks AraV, the regulator most likely remains bound to the *araS* promoter and other 'ara-box' containing DNA elements, thus repressing a subset of arabinose controlled genes. To the group of genes where 'ara-box' is present in the promoter region belongs *araS (Sso3061)* and *araD (Sso3124) (Brouns et al., 2006)*, which were down regulated in the deletion strain (Table 1). Importantly, in this scheme, the regulator does not interact directly with arabinose, but shuttles between an AraV bound and free state depending on the occupancy of the arabinose transporter with substrate.

Acknowledgements

Work of J.M.L. was supported by Ubbo-Emmius scholarship founded by University of Groningen. Work of S.V.A. was supported by a VIDI grant from NWO (the Dutch Science Organization).

Table.2. Strains and plasmids used.

Strain or plasmid	**Genotype**	**Derivative**
***Sulfolobus* strains**		
S. solfataricus PBL2025	*Δ(SSO3004–3050)*	PBL2000 (Worthington *et al.*, 2003)
S. solfataricus Δ*araV*	*araV::lacS*	PBL2025 (Schelert *et al.*, 2004)
S. solfataricus M16	*ΔlacS, ΔpyrEF*	*S. solfataricus* P1 (Jonuscheit *et al.*, 2003)
S. acidocaldarius MR31	*ΔpyrEF*	(Grogan and Hansen, 2003;Berkner *et al.*, 2007)
***E. coli* strains**		
Tuner + RIL	RIL plasmid produces archaeal tRNA	Novagen/ Stratagene
C43 (DE3)+ RIL	RIL plasmid produces archaeal tRNA	Miroux and Walker, 1996/ Stratagene
Stbl4™	Genetic markers allow stable cloning of direct repeat and retroviral sequences	Invitrogen
Plasmids		
pSVA146	*SSO3188, C-terminal his- and streptag*	pMZ1 (Zolghadr *et al.*, 2007)
pSVA147	*SSO3061, C-terminal his- and streptag*	pMZ1 (Zolghadr *et al.*, 2007)
pSVA148	*SSO3061, C-terminal histag*	pSA4 (Albers *et al.,* 2003)
pSVA149	*SSO3188, C-terminal histag*	pSA4 (Albers *et al.,* 2003)
pSVA154	*SSO3061, C-terminal his- and streptag*	pCMal (unpublished)
pSVA155	*SSO3188, C-terminal his- and streptag*	pCMal (unpublished)
pSVA167	*SSO3061, N-terminal histag*	pET15b (Novagen)
pSVA168	*SSO3188, N-terminal histag*	pET15b (Novagen)
pSVA169	*SSO3061, C-terminal streptag*	pSVA10 (Albers *et al.*, 2006)
pSVA170	*SSO3188, C-terminal streptag*	pSVA10 (Albers *et al.*, 2006)
pSVA171	*SSO3061, C-terminal streptag*	pMJ0503 (Jonuscheit *et al.*, 2003)
pSVA172	*SSO3188, C-terminal streptag*	pMJ0503 (Jonuscheit *et al.*, 2003)